Asymmetric cryo-EM structure of the canonical Allolevivirus Qβ reveals a single maturation protein and the genomic ssRNA in situ

Karl V. Gorzelnik1,1, Zicheng Cui2,1, Catrina A. Reed2,3, Joanita Jakana4, Ry Young5, and Junjie Zhang6,3

1Center for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843; 2National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030; 3Center for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843; and 4National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030

Edited by David DeRosier, Brandeis University, Waltham, MA, and approved August 5, 2016 (received for review June 21, 2016)

Single-stranded (ss) RNA viruses infect all domains of life. To date, for most ssRNA virions, only the structures of the capsids and their associated protein components have been resolved to high resolution. Qβ, an ssRNA phage specific for the conjugative F-pilus, has a T = 3 icosahedral lattice of coat proteins assembled around its 4,217 nucleotides of genomic RNA (gRNA). In the mature virion, the maturation protein, A2, binds to the gRNA and is required for adsorption to the F-pilus. Here, we report the cryo-electron microscopy (cryo-EM) structures of Qβ with and without symmetry applied. The icosahedral structure, at 3.7-Å resolution, resolves loops not previously seen in the published X-ray structure, whereas the asymmetric structure, at 7-Å resolution, reveals A2 and the gRNA. A2 contains a bundle of α-helices and replaces one dimer of coat proteins at a twofold axis. The helix bundle binds gRNA, causing denser packing of RNA in its proximity, which asymmetrically expands the surrounding coat protein shell to potentially facilitate RNA release during infection. We observe a fixed pattern of gRNA organization among all viral particles, with the major and minor grooves of RNA helices clearly visible. A single layer of RNA directly contacts every copy of the coat protein, with one-third of the interactions occurring at operator-like RNA hairpins. These RNA–coat interactions stabilize the tertiary structure of gRNA within the virion, which could further provide a roadmap for capsid assembly.

Significance

Single-stranded (ss) RNA viruses have ribonucleic acid as their genetic material and infect animals, plants, and bacteria. Here we used cryo-electron microscopy to reveal the genomic RNA (gRNA) of the ssRNA virus Qβ. The asymmetric gRNA adopts a single dominant structure in all virions and binds the capsid of Qβ at each coat protein. At the same time, we determined the structure of the maturation protein, A2, which functions both as the virion’s “tail” and its lysis protein. We see the gRNA is more ordered when interacting with A2. These results provide structural insights into gRNA packaging and host infection in ssRNA viruses.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the EMDataBank database, www.emdatabank.org (accession nos. EM-D-8253, EM-D-8254, and EM-D-8255). The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 5KIP).

See Commentary on page 11390.

*Center for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843; and National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030

1K.V.G. and Z.C. contributed equally to this work.

2Present address: Molecular and Cellular Medicine, College of Medicine, Texas A&M Health Science Center, College Station, TX 77843.

3To whom correspondence should be addressed. Email: junjiez@tamu.edu.

This article contains supporting information online at www.pnas.org/cgi/doi/10.1073/pnas.1609482113/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1609482113

PNAS | October 11, 2016 | vol. 113 | no. 41 | 11519–11524
the crystals (20, 21). The first glimpse of RNA within the capsid came from a low-resolution structure of MS2 using single-particle cryo-EM (27). On the basis of another single-particle cryo-EM structure with icosahedral symmetry applied, it was proposed that MS2 has two concentric shells of RNA within the phage capsid (28). Subsequent work, using cryo-electron tomography and subtomogram averaging, yielded a 39-Å resolution symmetry-free density map of MS2 (29), suggesting the gRNA adopts the same conformation for each virus particle within the capsid and the maturation protein replaces a C/C dimer of the coat proteins. However, at this low resolution, it was not clear how the maturation protein, which contributes ~1% of the molecular mass of the entire virion, interacts with the rest of the capsid or how the gRNA is organized inside the capsid shell.

In this study, we report the cryo-EM structures of the canonical *Allolevivirus* Qβ with and without symmetry applied, at 3.7- and 7-Å resolutions, respectively. Our structures reveal features never seen before for Qβ, such as the structure of A2, a symmetry deviation of the coat proteins, organization of gRNA, and the interactions between them. These results are discussed in terms of a model for viral assembly and gRNA release in Qβ.

**Results**

**Icosahedral Structure of Phage Qβ.** Ice-embedded Qβ phage particles were imaged using cryo-EM (SI Appendix, Fig. S2). A total of 712 image stacks were recorded on a direct detection camera in superresolution electron counting mode (30), yielding 51,815 high-quality phage particle images. Icosahedral symmetry was first applied to generate a structure of the coat protein shell at 3.7-Å resolution (SI Appendix, Fig. S3). Fig. 1 shows the icosahedral cryo-EM density map of Qβ with 180 copies of the 14-kDa coat protein. Each asymmetric unit within the capsid consists of three conformers: A, B, and C (Fig. 1A). Both types of dimers (A/B and C/C) are revealed. Protein α-helices and β-strands are clearly resolved throughout the cryo-EM density, with most of the bulky side-chains visible (Fig. 1B and C). The cryo-EM structure is consistent with the previously solved crystal structure of the icosahedral Qβ phage capsid. PDB coordinates of the crystal structure fit into the cryo-EM density without the need for major adjustments (SI Appendix, Fig. S4). Two loops within the coat protein of note are the EF-loop, which protrudes into the interior of the capsid, and the FG-loop, which forms small pores at both the fivefold and threefold vertices of the capsid (SI Appendix, Fig. S5) (20).

In our cryo-EM structure, the previously missing protein segment (residues 76–79) at the tip of the FG-loop of both conformers A and C is resolved (Fig. 1D), allowing us to build a more complete model for the capsid. The additional resolved residues are at the threefold vertices, with alternating FG-loops from conformers A and C. These residues form an ~15-Å pore, in agreement with its predicted size (20). Residues 76–79 of conformer B are resolved both in our cryo-EM reconstruction and the previous crystal structure, mostly because of the location at the fivefold vertices, which allows smaller pores (~8 Å in diameter) between neighboring subunits, restricting the movement of their FG-loops.

**Asymmetric Structure of Phage Qβ.** To determine the structures of the A2 protein and gRNA within the Qβ virion, we released the symmetry for our cryo-EM reconstruction. After 3D classification (Experimental Procedures), 12,975 particles were used to generate a symmetry-free reconstruction of Qβ at 7-Å resolution, with A2 and the gRNA clearly defined. The resolution for the asymmetric reconstruction of the coat proteins and A2 was further improved to 6.5 Å after masking out the RNA during map refinement (SI Appendix, Figs. S3 and S6). Figs. 2 and Movie S1 show the asymmetric cryo-EM structure of Qβ. The overall capsid stays similar to the icosahedral reconstruction of the coat proteins. The read-through domain of the A1 coat protein is not visible in our asymmetric reconstruction, possibly because of the flexible connector to the coat domain and/or the heterogeneity in the locations of A1 proteins among different virions. The most apparent difference in the capsid between our symmetric and asymmetric reconstructions is for A2 (colored hot pink in Fig. 2). A2, replaced one of the C/C dimers at a twofold axis (Fig. 2A), which leaves 178 copies of the coat protein in the virion. Much of the gRNA is clearly resolved, with major and minor grooves of the RNA helices discernible (SI Appendix, Fig. S7).

**Structure of the Maturation Protein A2.** The Qβ maturation protein, A2, consists of two structural regions positioned at an angle of 120° (Fig. 2). The top structural region, named the β-region, consists mostly of sheet-like densities from β-strands, whereas the bottom segment, termed the α-region, is mostly composed of rod-like densities from α-helices. Five α-helices were identified in the α-region and labeled α1 to α5, from the longest to shortest helix lengths (Fig. 2B), with α1 being ~60 amino acids. Sequence-based secondary structure analysis predicted several α-helices in A2 (SI Appendix, Fig. S8). Remarkably, one predicted α-helix (from residues 130–190) has a high confidence score and is consistent in length with α1. The α-region connects the gRNA through the capsid, the β-region, in contrast, is exposed on the surface of the phage with a tilt angle of 30° from the capsid shell (Fig. 2C). The dimers around A2 are numbered clockwise from 1 to 10 (Fig. 2D). The top halves of helices α1 and α2 interact with two A/B dimers (labeled 1 and 10) of the coat proteins. The bottom halves of helices α1 and α2, along with α3–α5, insert into the gRNA. Although most of the A2–gRNA interactions occur at the bottom of the α-helix bundle, one RNA stem-loop, labeled with an asterisk in Fig. 2, leans toward the top of the helix α5, forming an interaction between RNA with the middle of A2 (seen as the black dashed oval in Fig. 2B). Another RNA
Fig. 2. Asymmetric structure of the Qβ virus at 7-Å resolution. (A) Top view of Qβ with coat protein conformers A, B, and C colored light red, light green, and light blue, respectively. The maturation protein A2 is hot pink. (B) Segmented density of A2 showing the top β-region and bottom α-region. Five α-helices were modeled into A2. They are labeled from α1 to α5 according to their lengths, with the longest as α1. Black dashed oval indicates the location where one RNA stem-loop contacts. (C) Side view of Qβ rotated 90° from the top view, with A2 pointing up. The virion is sliced down the middle to reveal the interaction of A2 with the RNA (yellow). The RNA density is low-pass-filtered to 10-Å resolution. Densities of two RNA stem-loops, which are close to A2, are labeled with star and pound signs. (D) Zoom-in top view around A2 with the hot pink helix models shown for the α-helix bundle of α1 to α5. Ten coat protein dimers around A2 are labeled from 1 to 10 clockwise and colored salmon, green, and blue for conformers A, B, and C, respectively. Black arrows label the disruption of the interface between dimers 3 and 4. The yellow densities of two RNA stem-loops are labeled with star and pound signs as in C.

stem-loop, labeled with a pound sign in Fig. 2, displaces the FG-loop on conformer A of dimer 5 (SI Appendix, Fig. S9 and Movie S2). The presence of A2 and the gRNA forces the coat proteins to deviate from the perfect icosahedral arrangement seen in our symmetric reconstruction (SI Appendix, Fig. S9 and Movie S3), especially for the coat proteins around A2. Notably, the deviations are larger for dimers 1–5 compared with dimers 6–10. The distance between the protein backbones of opposite FG-loops increases from 20 Å for dimers 6–10 to 30 Å for dimers 1–5. The break in symmetry around A2 causes a fissure between dimers 3 and 4, disrupting the interactions that hold them together (Fig. 2D, SI Appendix, Fig. S9, and Movies S2 and S3), leaving more of an opening on the right toward which the β-region of A2 is angled. These disrupted coat protein interactions may allow for an easier detachment of A2, along with the gRNA, from the capsid upon infection. Furthermore, although A2 interacts with neighboring coat proteins, it does not make up for all of the interactions that would have taken place with the two coat proteins for which it has substituted. As a result, the weakened coat protein interactions around A2 may cause less obstruction during release of the gRNA from the capsid.

Organization of the gRNA. In the asymmetric structure, the gRNA exhibits well-defined secondary and tertiary structures within the ~30-Å-thick coat protein shell. Shown in Fig. 3 is a cutaway view of the asymmetric cryo-EM density map of Qβ oriented with A2 at the “North Pole” and its β-region, pointing out of the figure. The RNA elements (yellow density in Fig. 3A) show clear RNA helices and junctions. Only one shell of RNA lies beneath the capsid, with RNA further inside the virion showing distinct asymmetric tertiary structures. Several long RNA helices are observed to traverse the interior of the capsid (Movie S1). One striking feature is a 200-Å-long RNA helix (density fitted with a red RNA helix model in Fig. 3A) spanning the middle of the capsid close to its equator. One end of this long helix touches the inner surface of an A/B dimer that is slightly above the equator (marked with a purple oval in Fig. 3A and B); the other end of the long helix touches the inner surface of a C/C dimer on the equator at the opposite side (marked with an orange oval in Fig. 3A and C). In our asymmetric reconstruction, ~60% of the RNA density is located above the equator, and ~40% below. South of the equator, the local resolution of the RNA is lower compared with the RNA in contact with the capsid or proximal to A2 (SI Appendix, Fig. S6). This indicates that the RNA is more flexible in the interior of the southern hemisphere, possibly because of it being less packed in this region.

Interactions Between the Capsid and gRNA. The EF-loop of each coat protein thrusts inside the capsid, toward the RNA (Fig. 4A and SI Appendix, Fig. S5). Close inspection of the coat protein sequence of Qβ reveals that there are several positively charged residues at the tip of the EF-loop: residues 56–60 with the amino acid sequence SRNRK. These positively charged side-chains of Arginine and Lysine, at the tip of its EF-loop, may help anchor the coat proteins of Qβ to the negatively charged gRNA backbone during assembly. Moreover, it was previously proposed that, in many ssRNA viruses, the capsid proteins consist of highly basic semiflexible peptide arms, which provide nonspecific electrostatic interactions to control both the length and conformations of the gRNA (31).

We highlighted the tip of the EF-loops as spheres on our ribbon model of the coat proteins with colors salmon, green, and blue for conformers A, B, and C, respectively (Fig. 4A). Fig. 4B shows the density of gRNA and A2 with the locations of EF-loops labeled as dots in corresponding colors. After defining A2 as the North Pole, similar to that of Earth, we can unwrap the outermost spherical shell of the gRNA such that it would be like a map of the world. This is referred to as a Mercator projection, where the poles of the globe appear much larger in the projection than parts at the equator (Fig. 4C). What is readily seen is that each colored dot, representing EF-loops from the different-colored coat protein conformers, is immediately proximal to RNA, with most of the EF-loops appearing to be within the grooves of the RNA helices or at the tips of stem-loops. Interestingly, at some of the fivefold vertices, surrounded by the salmon ellipsoids (EF-loop of conformer A), the RNA also adopts a pseudo fivefold arrangement...
northern hemisphere, close to A2 (Fig. 5). Hairpin-coat protein interactions are asymmetrically distributed over the whole capsid in a Mercator projection, the operator-like preference of A/B dimers (23 of the 31 dimers). When visualized with both types of dimer, A/B and C/C (Fig. 5), with high binding affinity. The operator-like hairpins interact with the gRNA, suggesting they may share a similar mode of interaction with A2 in the North Pole and pointing out of the paper. Proteins and RNA are colored as in Fig. 2. The density of a long RNA helix, ∼200 Å in length, is docked with a red RNA duplex model and spans the two ends of the equator. Purple and orange ovals indicate the left and right of the helix, respectively. (B) Left outside view of the Qβ map, with the purple circle annotating the A/B dimer at the left end of the long helix in A. (C) Right outside view of the Qβ map, with the orange circle annotating the C/C dimer at the right end of the long helix in A. The black dashed lines, in B and C, label the equator.

Fig. 3. gRNA organization inside the Qβ capsid. (A) Cutaway front view of Qβ with A2 in the North Pole and pointing out of the paper. Proteins and RNA are colored as in Fig. 2. The density of a long RNA helix, ∼200 Å in length, is docked with a red RNA duplex model and spans the two ends of the equator. Purple and orange ovals indicate the left and right of the helix, respectively. (B) Left outside view of the Qβ map, with the purple circle annotating the A/B dimer at the left end of the long helix in A. (C) Right outside view of the Qβ map, with the orange circle annotating the C/C dimer at the right end of the long helix in A. The black dashed lines, in B and C, label the equator.

(Fig. 4C), which provides evidence that the structure of the outmost shell of the RNA depends on steric contributions from the capsid inner surface topography (32).

The interactions between coat proteins and gRNA in Qβ were first seen through a crystal structure of the complex between the operator and a coat protein dimer (24). We can fit this crystal structure (PDB ID code 4L8H) as a template into our asymmetric density map to look for interactions that are similar to the one between operator and coat protein dimers. To our surprise, 31 of the 89 coat protein dimers interact with operator-like hairpins, suggesting they may share a similar mode of interaction with high binding affinity. The operator-like hairpins interact with both types of dimer, A/B and C/C (Fig. 5A and B), with a preference of A/B dimers (23 of the 31 dimers). When visualized over the whole capsid in a Mercator projection, the operator-like hairpin-coat protein interactions are asymmetrically distributed within the capsid, mostly between longitudes 90° and 180° in the northern hemisphere, close to A2 (Fig. 5C and D). The coat protein dimers also interact with the gRNA in other locations, such as in the middle of RNA helices or on RNA junctions (SI Appendix, Fig. S10).

Discussion
Role of A2 as the Maturation and Lysis Protein. The results presented here may have implications for the mechanism of host infection and of gRNA packaging in ssRNA phages. The Qβ gRNA is surrounded by a coat protein shell, whose symmetry is broken by the maturation protein, A2. A2 replaces a C/C dimer with its β-region tilting at an angle of ∼30° from the virus surface, just barely rising above the capsid. As in all Leviviridae, the maturation protein A2 is required for infection. The β-region of A2, outside the capsid, may carry out the role of binding to the receptor, the F-pilus, whereas the α-region binds to the gRNA and hangs on until the RNA is inside the cell. In Qβ and Qβ-like phages, the portion of A2 outside the capsid has the expanded role of causing lysis. At the time of lysis, it has been shown that A2, either free in the cytosol or as a single protein on the virion, can both bind and inhibit MurA (33), a universally conserved enzyme that catalyzes the first committed step in cell wall biosynthesis. From our asymmetric structure of Qβ, the MurA-binding site of A2 is likely to be exposed and around the β-region.

Structural Implication for gRNA Release. We see that A2 coordinates denser packing of gRNA around it, locally expanding the capsid and causing a fissure between coat dimers (SI Appendix, Figs. S9 and S3). Upon attachment to the F-pilus, the β-region of A2 may have weakened interactions with the neighboring coat proteins, easing the detachment of A2-bound gRNA. Upon infection, most of the RNA tertiary interactions, as well as the gRNA–coat interactions, need to be abolished for the gRNA to exit through the narrow slot (∼60 × 40 Å), left on the capsid after A2 is detached. As ssRNA viruses of Leviviridae do not have a pressurized capsid, the gRNA cannot be mechanically injected into the host from the capsid. The retractile character of the F-pilus may provide some or all of the driving force to pull the A2-bound gRNA out of the capsid. Notably, the loose interactions of A2 with the coat proteins and the disrupted interfaces between coat protein dimers around the exit may facilitate the gRNA release. The fissure between coat protein dimers 3 and 4 may further

Fig. 4. EF-loops anchor the coat proteins onto the outmost shell of gRNA. (A) Interaction between the gRNA and an asymmetric unit of the coat protein with the residues at tip of the EF-loop shown as sphere models. The density of gRNA is in yellow, whereas the models of the coat proteins are colored salmon, green, and blue for conformers A, B, and C, respectively. (B, Left) A2 (hot pink) and Qβ gRNA (yellow) are viewed with A2 as the North Pole (labeled “N”). Tips of the EF-loops are presented as colored dots (with color conventions as in A) around the Qβ gRNA globe. (Right) Earth in the same orientation. (C) Mercator projection of the outmost shell of the Qβ gRNA (yellow) and the locations of EF-loops (salmon, green, and blue ellipsoids).
and shown as the North Pole (labeled capsid density with coat protein dimers colored as in Gorzelnik et al. PNAS (9). The final protein concentration of Q capsid may enable the gRNA to sample different conformations to is emptied. The loose packing in the southern hemisphere of the expand during gRNA release and come together when the capsid
Fig. 5. coat proteins are produced, the rate-limiting step in virion pro-
interactions with neighboring coat protein dimers, which in turn possibly presenting a roadmap to sequentially recruit more coat
and blue for their conformers A, B, and C, respectively. Dimers not interacting with operator-like hairpins are in gray. The α-helix bundle of A is in hot pink and shown as the North Pole (labeled “N”). (D) Mercator projection of the capsid density with coat protein dimers colored as in C.

Structural Mechanism for Capsid Assembly. For Qβ and other ssRNA viruses, there is clearly an evolutionary mandate for efficient and accurate virion assembly (34). After the gRNA is replicated and coat proteins are produced, the rate-limiting step in virion production is encapsidation of the RNA (35). It was first proposed that the assembly pathway was initiated when the first coat protein dimer bound the operator hairpin, serving as a nucleation event for other coat protein dimers to “walk” down a “Hamiltonian path” binding neighboring RNAs until the virion is complete (36). To accommodate this perspective, it was later proposed that there are many sites within the gRNA, termed “packaging signals,” that can interact with the coat protein (6). These packaging signals were thought to be related to the consensus sequence of the operator. In our re-
construction of Qβ, we see the outmost layer of the gRNA coming in contact with almost every coat protein, presenting 31 operator-
lke hairpins for the binding of coat protein dimers. Many of the gRNA secondary structures involve an extensive array of long-range tertiary interactions, which have been shown to play roles in gene regulation and capsid assembly (37). These tertiary interactions may contract the gRNA to a rigid structure in which the operator-like hairpins, which may be far apart in the genome, are located near each other. As seen in our structure, the operator-like hairpins are asymmetrically distributed on the outmost shell of the gRNA, possibly presenting a roadmap to sequentially recruit more coat protein dimers. The recruited coat protein dimers strengthen their interactions with neighboring coat protein dimers, which in turn stabilize the gRNA underneath. Such a mutual chaperoning of both RNA and coat protein conformations (38) may facilitate the fast and accurate assembly of these viruses in vivo.

Experimental Procedures
Sample and Grid Preparation. Phage Qβ was purified as previously described (9). The final protein concentration of Qβ was 5 mg/mL, ~10^12 plaque-forming units. For imaging, 3 μL of the sample solution was first applied to a C-Flat 1.2/1.3 400 mesh holey carbon grid at 20 °C with 100% relative humidity and vitrified using a Vitrobot Mark III (FEI Company).

Cryo-EM Imaging. The thin-ice areas that showed clearly visible and mono-
dispersed particles were imaged under a JEM3200FSC cryo-electron micro-
scope with a field emission gun (JEOL) operated at 300 kV. Data were collected using the manual mode of SerialEM (39) on a Gatan K2 Summit direct detection camera (Gatan) in the supersolution electron counting mode. A nominal magnification of 30,000× was used, yielding a subpixel size of 0.6 Å. The beam tilt intensity was adjusted to a dose rate of 7.2 e^-/Å^2 per pixel on the camera. A 50-frame movie stack was recorded for each picture, with 0.2 s per frame, for a total exposure time of 10 s. Each frame in the movie stack had an exposure dose of 1 e^-/Å^2 per frame. An in-column energy filter was used with a slit width of 29 eV.

Data Preprocessing. A total of 712 superresolution movie stacks were col-
clected and first binned by 2 to yield a pixel size of 1.2 Å. Using Unblur (40), these stacks were aligned, filtered according to electron dose, and summed to generate two sets of sum images, with one set from frames 1–30 and the other set from frames 2–12, respectively.

These sum images were visually screened, and 582 of each set with strong power spectra were selected for further processing. Contrast transfer func-
tions of the micrographs were estimated using CTFfind (41). Batchboxer in EMAN (42) was first used to automatically pick all the particles from these 582 sum images with a box size of 320 × 320 pixels, yielding 111,507 parti-
cles. The automatically picked particles were then shrunk by eight times to a pixel size of 9.6 Å and screened for high-contrast particles for three rounds of the reference-free 2D classification in Relion (43), with each round having 18, 16, and 25 iterations, respectively. The coordinates of the 53,999 screened particles were then imported into the ezboxer.py in EMAN2 (44) to manually screen for missed false positives, leaving only 51,815 particles.

Icosahedral Refinement for the Coat Proteins. The 51,815 clean particles were then refined in Relion, with the icosahedral symmetry applied. Initial re-
finement was run using the sum images from frame 1–30 (total dose of 31 e^-/Å^2 with dose filter on in Unblur). The starting map was generated from EMAN2 and low-pass-filtered to 60-Å resolution. The refinement first yielded a map at 4.0-Å resolution. This was followed by a continuing map refinement with sum images of only frame 2–12 (total dose of 11 e^-/Å^2 with dose filter on) and the local contrast transfer functions for each particle reestimated using Gctf (45). The final refined icosahedral map yielded a resolution of 3.7 Å.

Asymmetric Refinement for the Entire Virion. The same 51,815 clean particles were binned by 2 to a pixel size of 2.4 Å for the asymmetric reconstruction of the entire virion. The icosahedral map of the coat proteins was low-pass-
filtered to 60 Å as an initial model. Unsupervised 3D classification in Relion was performed with C1 symmetry for 9,675 sum particles and 12,975 particles (46), yielding 111,507 parti-
cles, which showed clear density of the gRNA inside the capsid. These 12,975 particles were further refined in Relion with C1 symmetry to yield a map at 7.0-Å resolution, where the protein secondary structures were clearly visible. To further improve the resolution of the coat proteins and A, a mask was applied to focus the refinement on the outer shell of the particle, which yielded a resolution of 6.5 Å. Further refinement of the asymmetric structure using all of the 51,815 particles did not improve the resolvability of the map, potentially because the 12,975 particles, selected from the 3D classification in Relion, have the best contrast and are most homogeneous in conformations.

Resolution Estimation. The overall resolutions of all these reconstructed maps were assessed using the gold-standard criterion of Fourier Shell Correlation, with a cutoff at 0.143, between two half maps from two independent half-
sets of data (46). Local resolutions were estimated using Blocres (47).

Modeling, Map Segmentation, and Visualization. To build the model for the coat protein from the icosahedral reconstruction, the crystal structure (PDB ID code 1QBE) was used as the initial model. The missing loop in the crystal structure was built in COOT (48) and then further refined in the real-space refinement in Phenix (49). The refined model showed good model geometry from Molprobity (50) (SI Appendix, Table S1). The Fourier Shell Correlation between the icosahedral map and the refined PDB model agreed at 3.9-Å resolution at 0.5 threshold (SI Appendix, Fig. S3). This more complete model for the coat protein was then used to fit into the asymmetric map of the Qβ virion to guide the segmentation of the gRNA and A2 protein with Segger (51) as a plugin in University of California, San Francisco (UCSF) Chimera (52).

Gorzelniak et al. 11524 | www.pnas.org/cgi/doi/10.1073/pnas.1609482113

ACKNOWLEDGMENTS. We thank Jeng-Yih Chang for performing the molecular dynamics flexible fitting of the long RNA helix model. We thank the Microscopy and Imaging Center at Texas A&M University for providing instrumentation for the initial screening of the phage particles; the Texas A&M High Performance Research Computing Center for providing the computational resources for the data processing; and the National Center for Macromolecular Imaging (NCMI) at Baylor College of Medicine for data collection. NCMI is supported by NIH Grants P41GM103832 and U24GM116787. J.Z. is supported by start-up funding from the Department of Biochemistry and Biophysics at Texas A&M University. R.Y. is supported by Public Health Service Grant GM72099. J.Z. and R.Y. acknowledge the support of the Center for Phage Technology, jointly sponsored by Texas AgLife and Texas A&M University. K.G.V. is partly supported by the Welch Foundation Grant A-1863 (to J.Z.).
Supporting Information
Gorzelnik et al. 10.1073/pnas.1609482113

Movie S1. Asymmetric structure of Qβ. Coat proteins are in salmon (conformer A), green (conformer B), and blue (conformer C), respectively. A₂ is in hot pink. RNA is in yellow and low-pass–filtered to 10-Å resolution.

Movie S2. Symmetry breaking in the coat proteins around A₂.
Movie S3. The deviations of the coat protein dimers in the asymmetric structure from their locations in the icosahedral structure.

Movie S3

Other Supporting Information Files

SI Appendix (PDF)
SI Appendix:

Asymmetric cryo-EM structure of the canonical Allolevivirus Qβ reveals a single maturation protein and the genomic ssRNA in situ

Karl V. Gorzelnik\textsuperscript{a,1}, Zhicheng Cui\textsuperscript{a,1}, Catrina A. Reed\textsuperscript{a,2}, Joanita Jakana\textsuperscript{b}, Ry Young\textsuperscript{a}, and Junjie Zhang\textsuperscript{a,3}

\textsuperscript{a} Center for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843

\textsuperscript{b} National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030

Movie Legends

Movie S1. Asymmetric structure of Qβ. Coat proteins are in salmon (conformer A), green (conformer B) and blue (conformer C), respectively. A\textsubscript{2} is in hot pink. RNA is in yellow and low-pass filtered to 10Å resolution.

Movie S2. Symmetry breaking in the coat proteins around A\textsubscript{2}.

Movie S3. The deviations of the coat protein dimers in the asymmetric structure from their locations in the icosahedral structure.
Figure S1. Genomic maps of *Leviviridae* MS2 (top; Levivirus) and Qβ (bottom; Allolevivirus). The maturation proteins are highlighted in hot pink. The lysis protein of MS2 is highlighted in purple. A₁ of Qβ is labeled as dashed box. Yellow stars indicate the operators near the start of their respective replicase genes.
Figure S2. Drift-corrected micrograph of Qβ (left) and its power spectrum (right).
Figure S3. Gold-standard Fourier shell correlation (FSC) for the cryo-EM reconstructions of Qβ with and without symmetry and between the icosahedral map and the refined PDB model.
Figure S4. Fitting of the crystal structure (PDB ID 1QBE) into our icosahedral cryo-EM density map at 3.7Å resolution.
Figure S5. PDB model of Qβ coat protein showing the location of EF-loop and FG-loop. The protein backbone is colored blue to red from N-terminus to C-terminus.
Figure S6. Local resolution of the asymmetric reconstruction of Qβ from outside (left) and cutaway (right) views. The yellow dashed circle encloses the gRNA densities.
Figure S7. Segmented density map of genomic RNA showing major and minor grooves.
Figure S8. Secondary structure prediction for the A₂ protein from I-TASSER, with the first line as peptide sequence, second line as predicted secondary structures and the third line as the confidence score of the predicted secondary structures. The larger the confidence score, the higher the probability of the predicted secondary structures.
Figure S9. Deviation of the coat protein dimers’ position in the asymmetric structure from their positions in the icosahedral reconstruction. (A) Zoom-in top view of the capsid proteins in the asymmetric structures (salmon, green and blue ribbon models) overlaid with their positions in the icosahedral structure (grey ribbon models). The α-region of A₂ is shown as a five-helix bundle (α1-5, hot pink ribbon models). Two RNA stem-loops (yellow density) are labeled with the star and pound symbols, respectively. Black arrows label the disruption of the interface between two dimers. (B) Zoom-out top view of asymmetric Qβ capsid with the black arrows indicating the direction and amount of deviation for each coat protein dimer from its position in the icosahedral structure. The length of each arrow equals 10 times the deviation between the corresponding dimer’s centers-of-mass in the icosahedral and asymmetric capsids. (C) The side view with viewing angle rotated 90° from the top view.
Figure S10. Examples of coat protein dimers (grey ribbon models) interacting with gRNA (yellow densities), which do not have an operator-like hairpin fold. (A) Side view and top view of one coat protein dimer interacting with the middle of an RNA helix. (B) Side view and top view of one coat protein dimer interacting with an RNA junction.
Table S1. Molprobity model statistics for the coat proteins built from the icosahedral map.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All-atom Clashscore</td>
<td>10.42</td>
</tr>
<tr>
<td>Poor rotamers</td>
<td>0.00%</td>
</tr>
<tr>
<td>Favored rotamers</td>
<td>98.06%</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td>0.00%</td>
</tr>
<tr>
<td>Ramachandran favored</td>
<td>83.08%</td>
</tr>
<tr>
<td>MolProbity score</td>
<td>2.23</td>
</tr>
<tr>
<td>Cβ deviations &gt;0.25Å</td>
<td>0.00%</td>
</tr>
<tr>
<td>Bad bonds:</td>
<td>0.00%</td>
</tr>
<tr>
<td>Bad angles:</td>
<td>0.07%</td>
</tr>
</tbody>
</table>